

Single Chain Force Spectroscopy: Sequence Dependence

Namkyung Lee, Thomas A. Vilgis

Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany

(February 1, 2008)

We study the elastic properties of a single A/B copolymer chain with a specific sequence. We predict a rich structure in the force extension relations which can be addressed to the sequence. The variational method is introduced to probe local minima on the path of stretching and releasing. At given force, we find multiple configurations which are separated by energy barriers. A collapsed globular configuration consists of several domains which unravel cooperatively. Upon stretching, unfolding path shows stepwise pattern corresponding to the unfolding of each domain. While releasing, several cores can be created simultaneously in the middle of the chain resulting in a different path of collapse.

I. INTRODUCTION

Recent advances in nano manipulation allow direct access to single molecule elasticity using Atomic Force Microscopy (AFM) [1–6]. AFM seems to be an excellent technique to study interactions in single chains by use of their elastic response. As a result of interplay between entropy, interactions between monomers and environments (such as solvents, other chains, external fields) the resulting force-extension relations allow deep insight into the nature of the interactions and conformations. In some polymer systems (including biopolymers and proteins), intrachain self assembly produces secondary or tertiary structures and the elastic responses reflect this structural hierarchy [7–9]. Furthermore, mechanically induced folding/unfolding leads a single molecule to undergo structural transitions following well defined specific paths without being frustrated. Especially in the case of proteins there is considerable hope to gain insight into folding mechanisms and design.

In the present letter we study the conformation of random A/B copolymers and their responses to external forces. As the response reveals information of the sequence and structure of the chain, this allows for “single chain force spectroscopy”. We use variational methods to provide some understanding of the elasticity of polymer with specific sequence. The present model corresponds also to the most simple “protein model” [10] when a polymer chain consists of more than one constituents (here A and B monomers) which are arranged in a specific sequence along the backbone. The sequence carries relevant information since according to the arrangement of the monomers certain conformations are more preferred than others.

Although the difficult nature of the problem suggests a detailed mathematical analysis (which is partly found below), some physical scaling ideas help to understand the problem. Blob pictures and estimates of relevant length scales provide an instructive physical picture for homopolymer deformation in poor solvent condition. Here relevant length scale is the thermal blob size ξ_T . Each blob carries energy of $k_B T$. A sharp structural transition occurs when the Pincus blob size $\xi_P (\equiv k_B T/f)$ reaches to this thermal blob size $\xi_P = \xi_T$.

Experimentally, two scenarios are available: either the displacement or the force are controlled [1–6]. When the displacement is imposed and the force is measured, the unfolding of sequence of domains leads to repeating “saw tooth” pattern in force extension curve [1,2]. In constant force measurement, folding-unfolding transition appears at characteristic force, which is shown as “plateau” [6]. In the present study we minimize the free energy at given force (constant force measurement). As we will show later, we find multiple minima at a given force. These metastable states are created by the heterogeneity of the sequence. Some of these configurations are visited depending on the history of folding/unfolding.

II. ANISOTROPIC VARIATIONAL METHOD

Formally we consider a polymer chain consisting of N monomers of size b under tension which is described by

$$H = H_c + H_i + f \cdot (\vec{r}_N - \vec{r}_1). \quad (2.1)$$

The first term corresponds to the elastic and connectivity properties of the polymer chain and is approximated by the Wiener measure [11]. The second term contains all interactions between monomers, especially, two and three

body interactions of the virial expansion are included here. The three body term is essential since depending on the solvent quality certain monomers may attract each other, which results in a (partial) collapse of the chain. The two body interactions can be attractive or repulsive depending on the type of the monomer pairs. The density of the collapsed globule is determined by balance between the three body repulsions and two body attractions. The last term is the elastic energy term which represents the external force field. The variational free energy is defined as usual by $F_V \equiv \langle H - H_t \rangle_0 + F_0$ where $\langle \cdots \rangle_0$ stands for the average over the variational probability distribution:

$$P_V(\vec{r}_1, \dots, \vec{r}_N) = Z_V^{-1} \exp\{-H_t(\vec{r}_1, \dots, \vec{r}_N)\}, \quad (2.2)$$

where Z_V is the normalization constant satisfying $\int P_V = 1$ and $F_0 = -k_B T \log Z_V$.

The presence of an external force breaks the spherical symmetry. We need, therefore, to distinguish the deformation in parallel and perpendicular direction with respect to the external force. In terms of two components in correlation function, we choose consequently the trial Hamiltonian H_t as follows

$$H_t(\vec{r}_1, \dots, \vec{r}_N)/k_B T = \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^N [G_{\parallel}^{-1}(j, l)(\vec{r}_{\parallel}^j \cdot \vec{r}_{\parallel}^l) + G_{\perp}^{-1}(j, l)(\vec{r}_{\perp}^j \cdot \vec{r}_{\perp}^l)], \quad (2.3)$$

where $G_{\parallel}(j, l)$, $G_{\perp}(j, l)$ are correlation functions between $j - l$ monomers in parallel and perpendicular direction with respect to the external force. By minimizing the trial free energy function with respect to G_{\perp} , G_{\parallel} , we obtain two sets of self-consistent equations. These self-consistent equations for both components are coupled, in contrast to systems with spherical symmetry as studied previously [12–14].

The square of the distance $b(i, j)$ between two monomers i and j is related to $G_{\parallel}(i, j)$ by

$$b_{\parallel}(i, j) = G_{\parallel}(i, i) + G_{\parallel}(j, j) - G_{\parallel}(i, j) - G_{\parallel}(j, i). \quad (2.4)$$

Therefore, the square of the mean end to end distance $\langle R^2(1, N) \rangle$ is determined by the quantity $b_{\parallel}(1, N)$. The detailed derivation and solution of the self-consistent equation is shown elsewhere [15]. This procedure corresponds to the self-consistent one-loop approximation [16].

III. STRETCHING A HOMOPOLYMER IN A POOR SOLVENT

As a test run we apply the variational calculation to calculate the elastic response of a homopolymer in a poor solvent. This problem has been studied theoretically (scaling [17], MC simulation [18]), and experimentally [19].

Using the present variational method, we confirmed the scaling picture of force-extension relations of homopolymer. The position of the plateau in each curve captures the characteristic force f_c at each solvent condition and is in agreement with scaling picture. When we follow the local minima, we observe a hysteresis effect at the boundary of the first order phase transition. Around f_c , we access to local minima with different extensions. This means that chain configuration becomes metastable. At the transition point, tadpole and ellipsoidal conformations coexist.

IV. STRETCHING A HETEROPOLYMER IN A POOR SOLVENT

Now we consider a random A/B polymer chain of length N . The interaction Hamiltonian for the pairwise interaction is given by

$$H_i = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N v_{ij} \delta(\vec{r}_i - \vec{r}_j), \quad (4.1)$$

where the two body interaction possesses a random nature, i.e., the interactions are chosen as a random $N \times N$ matrix [20–22]:

$$v_{ij} = v_0 + [\alpha(\sigma_i + \sigma_j) + \chi(\sigma_i \sigma_j)]. \quad (4.2)$$

The sequence of monomers is described by variables σ_i . $\sigma_i = -1$ if monomer i is of type A (say, hydrophobic) and $\sigma_i = 1$ if it is of type B (hydrophilic). To be more specific, a Flory like mixing parameter $\chi = (v_{AA} + v_{BB})/2 - v_{AB}$ is negative when similar monomers attract each other. The second virial coefficient $v_0 < 0$ provides that average attraction between monomers. Obviously, the results will depend on the specific sequence of type A and type B

monomers. For technical reasons we assume that the average solvent quality is poor compared to the fluctuation due to randomness ($|v_0| < \chi$), therefore the thermodynamic ground state is always a globular state. Before a further discussion on the force extension properties for given sequences, we consider the physical effects of the randomness in our model to see what we can expect.

V. STATISTICS

To get a physical understanding for the problem we average over the disorder and assume a Gaussian distribution for the disorder part in the two body interaction, i.e., $P(\sigma_i) = e^{-\sigma_i^2/2\delta}$. We obtain the first order correction in the effective binary second virial coefficient between arbitrary monomer pairs $l - m$: $\{v^{\text{eff}}\}_{lm} = v_0 - \alpha^2/\chi$ for $l \neq m$.

If the interactions between the similar monomers are equal ($v_{AA} = v_{BB}$) then we have $\alpha = 0$. We expect no effective change in the effective solvent quality. In our consideration, we assume that $\alpha \sim \chi$. The Flory mixing parameter χ represents also the strength of disorder. Each thermal blob size within the collapse globular phase will be represented by the effective second virial coefficient v_{eff} instead of v_0 , which leads to $\xi_T^{\text{random}} = b^4/v^{\text{eff}}$. The average thermal blob size is $\xi_T^{\text{random}} = \xi_T^{\text{homo}}(1 - \alpha/|v_0|)$. The mean square average of v^{eff} is $\sim \alpha/\sqrt{N}$. For small disorder, for each sequence, the structural transition occurs at large scale. The variance in characteristic force f_c reflects the size of disorder ($\sim \alpha\sqrt{N}$). Although the mean net charge is zero, a typical random copolymer of size N has excess charge of order $\pm\sqrt{N}$. For some sequences, there might be an excess charge $\sum \sigma_i \neq 0$. The characteristic force f_c is changing accordingly due to the shift in an effective solvent quality.

Depending on the mean square fluctuation, we divide the force-extension curve into three distinctive regime. For small extension ($\xi_P > \xi_T^{\text{largest}}$, $f < k_B T/\xi_T^{\text{largest}}$) a collapsed globule deforms to ellipsoidal conformation. For large extension ($\xi_P < \xi_T^{\text{smallest}}$, $f > k_B T/\xi_T^{\text{smallest}}$), all thermal blobs are linearly aligned. No conformational degree of freedom is left for rearrangement in thermal blobs. The blob size is then simply equal to Pincus blob size. In intermediate regime ($\xi_T^{\text{smallest}} < \xi_P < \xi_T^{\text{largest}}$), we expect fluctuations in $z(f)$ to depend on the system size and external force.

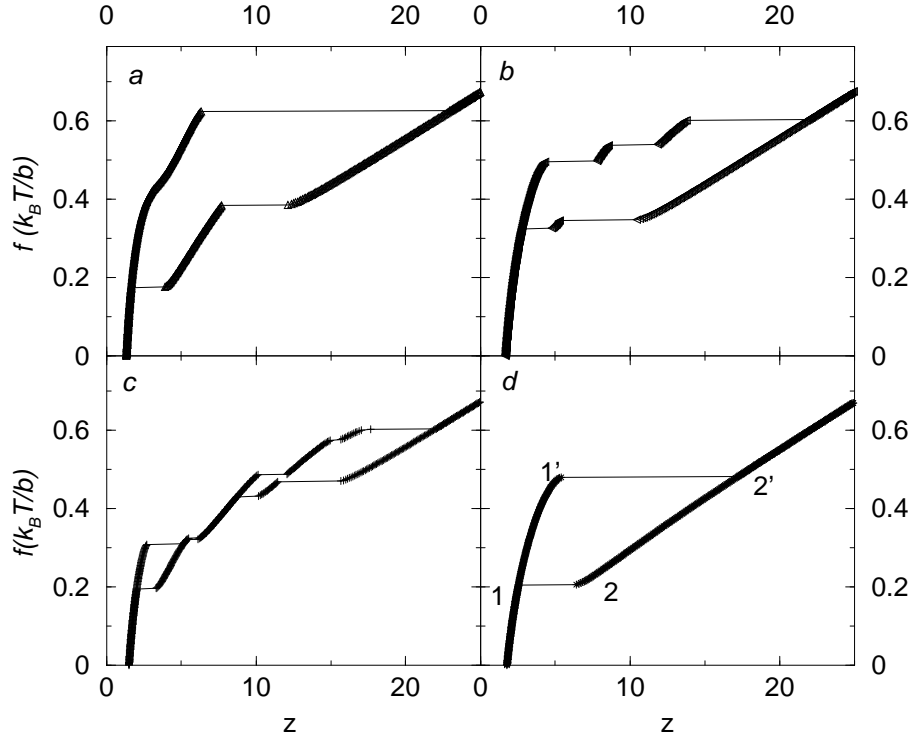


FIG. 1. Typical force-extension curve which show multiple steps for four different sequences with $\chi/|v_0| = 0.5$. The hysteresis during stretch-release cycle is shown at each curve. The sequence of each polymer is

- a) AAAAB BBBAB BBBB AABBA BABAA BAABA BAAAB BABAA
- b) BABAA BABBB BBAAA ABBAB BAAAB ABBAA ABBAB BBAAA
- c) ABAAB BAAAA AAAAB BAABB BBAAB BBBAB ABBBB BABAA
- d) AABAB AABBB AABAB BABBB AAABB BABAB ABBAA BABBA

VI. FORMATION OF DOMAINS: BLOCKS

Due to the heterogeneity of the sequences, in average, the structural transitions are smoothened. However, unfolding of individual chains shows unique elastic response reflecting the information in the interactions along the sequence. Within a collapsed globular phase, a local sequence of “n” monomers has an excess charge with typical fluctuations of size \sqrt{n} . This creates more hydrophobic (type A dominant) cores. Their sizes are determined by the disorder (in local sequence) and their positions along the chain are random. In fig. 1, we show typical examples for the force-extension relationships which show multiple steps for four different types of sequences with $\chi/|v_0| = 0.5$.

Multiple plateaus indicate that the force induced unfolding from the globular to open-string structure occur via intermediate structural forms. Each plateau in force-extension relation corresponds to the unfolding of a part of the chain cooperatively which defines a “domain”. The structural changes are achieved by abrupt unraveling of the corresponding domain. It is clearly visible that the curves depend on the sequence. For example, curve d) in fig. 1 corresponds to the most “alternating” sequence (alternating copolymer). Therefore, the overall poor solvent dominates, and the conformation is a single globule. In the situations of curves a) and d), several A and B blocks are separated by “non-blocky” units. These allow for a more structured phase space, the force-extension relation is richer. Moreover sequence c) has a net charge 2 (zero charge for a,b and d).

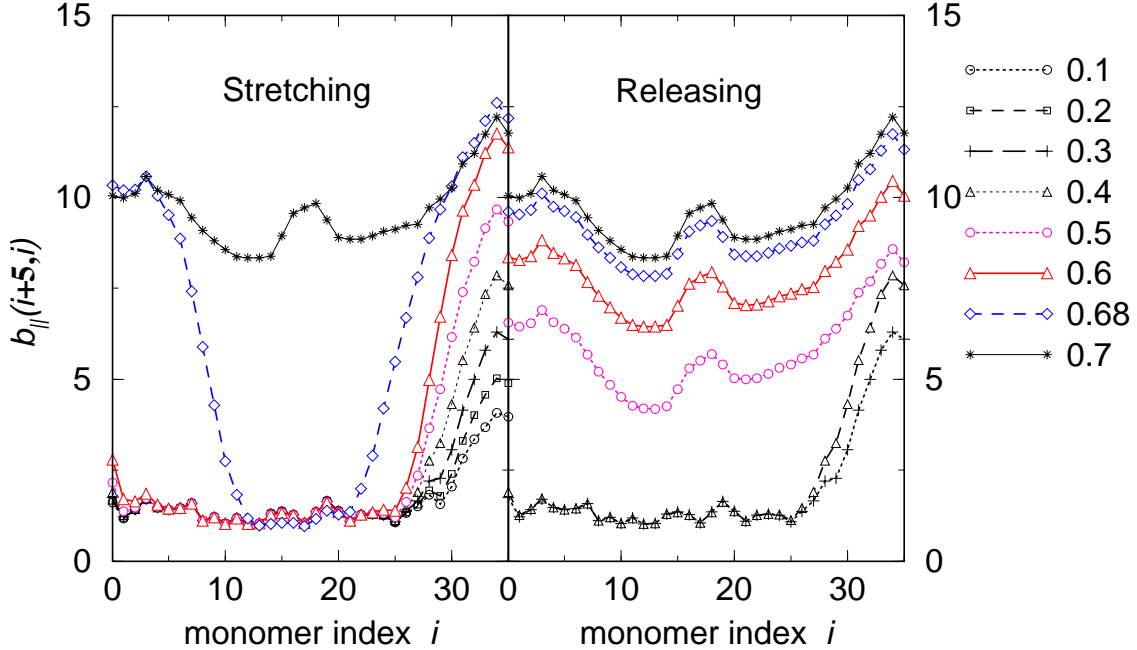


FIG. 2. The inverse correlation function $b_{||}(i+5, i)$ during the stretch-release cycle. The chain conformation choose different paths for folding and unfolding procedure. The correlation $b_{||}(i, j)$ of some part of chain increase/decrease cooperatively. Different symbols indicate applied force in units of $k_B T$.

In fig. 2 the correlation function of a specific sequence (c in fig. 1) is shown during the stretch-release cycle. The regions of small values of $b_{||}(i, j)$ represents the globular phase. As the force increases, a part of the chain is pulled out from one end leading to a tadpole like conformation. Upon further increase of the force, the remainders in the globular phases are reduced. Then, the unfolding of the globule occurs in stepwise patterns, which have also been predicted for polyelectrolyte necklace chains [23]. Different configurations are visited while the chain releases. The more hydrophobic parts assemble together at the beginning of the release, the necklace like conformations are formed. These clusters are growing and finally merge to a single globule. The fig. 2 also shows that upon releasing the structural change from necklace to globule occurs at smaller characteristic force than upon stretching. The hysteresis on force extension curves is determined by the energy barriers between two coexisting configurations following a force induced path. The time for relaxation is exponentially increasing with respect to the energy barrier ΔE^b between the two states, $t_r \sim \exp(\Delta E^b/k_B T)$. These local minima created by quenched disorder is stable with respect to the fluctuation [24]. The hysteresis appears in the experimental situation with finite pulling speed [2]. It has been already

seen that for pulling speeds faster than t_r , the rearrangement of monomers at given displacement is not followed by unfolding of each domain. However, a complete analysis of the dynamical aspects and the influence of the pulling velocity will be given in a separate paper [15].

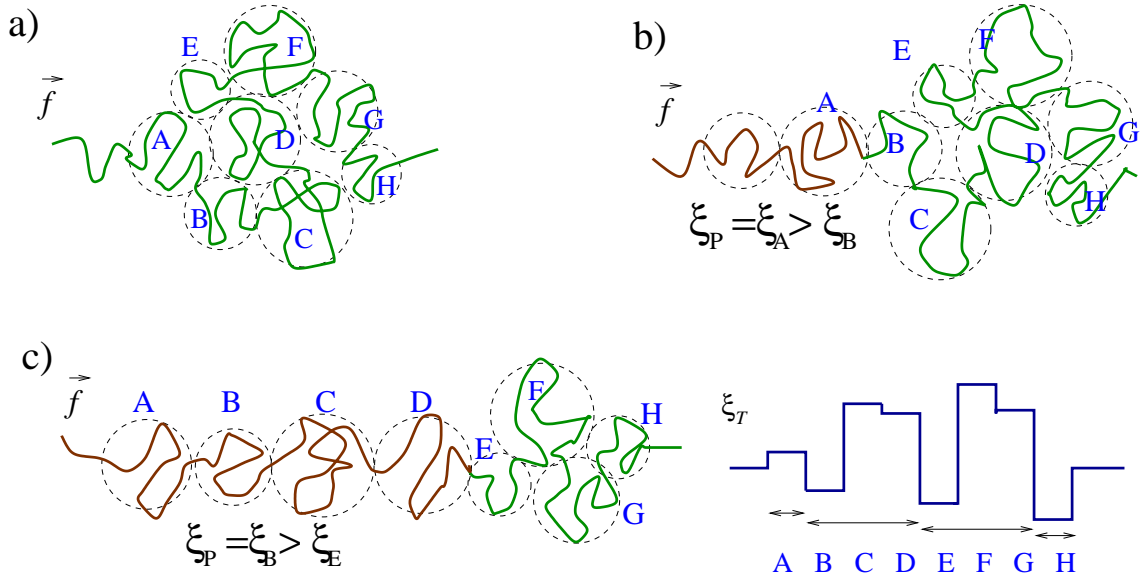


FIG. 3. Blob interpretation of stretching a random copolymer

The stretching without relaxation is illustrated in fig. 3. The external force propagates from ends to the center of the chain. The extension is controlled by smaller thermal blob size along the path of force propagation. Variations in thermal blob size ξ_i due to the disorder along the chain are expected. Each blob carries the thermal energy $k_B T$ and the variation in size of blob comes from the variation in local sequence. Applying external force unravels each thermal blob one after another when the force reaches the order of $\xi_P = \xi_i$ ($i=A, B, C, \dots$). The force needed to pull out blob B is large enough to pull blobs C and D (since $\xi_C, \xi_D < \xi_B$). Therefore, as soon as blob B is released, there is no further increase of the force required to pull C and D. This means that the B-C-D blobs respond to the external force cooperatively. In similar way, E-F-G form another domain. After being pulled out, the size of each blob should be rescaled to the size of Pincus blob at given force.

In thermodynamic equilibrium at given force f_0 , more than one configuration with different extension can be accessed with some probability. (For example, 1 and 2 in fig. 1 d). The ratio between the average life time of two configurations is $P(E_1)/P(E_2)$. This is related to the free energy difference $\Delta F = -k_B T \ln(P(E_1)/P(E_2))$. From the force extension curve, the free energy difference can be easily obtained: $\Delta F = F_1 - F_2 = f_0 \delta z_{1-2}$. Estimation of energy barrier along the path 1-1'-2'-2 (shown in Fig. 2 d) is $\Delta E_{1-2}^b = \int_1^{1'} f(z) dz$.

VII. CONCLUSION

We investigated the elastic properties of a single chain with arbitrary sequences using variational method. The essential property governing the force-extension curve is domain-wise unfolding from globular to open string conformation. The characteristic force related to each domain is captured as a “plateau” in the force-extension curve. In this variational approach, other types of interactions (including long range interactions) and other properties of the chain (e.g., stiffness) can be easily incorporated. When long range interactions are dominant, domains becomes correlated. This effect will smoothen out the sharp transition due to the unfolding of each independent domain. We discuss the role of long range interactions separately [15]. Proteins could be also treated in similar way by assigning special sequence using binary interaction matrix and by adding more types of monomers which may correspond to the amino acids. The accuracy of the method is tested in comparison with Monte Carlo simulation [25].

ACKNOWLEDGMENTS

We thank V. Rostiashvili, S. P. Obukhov and A. Johner for thoughtful discussion. N. L. would like to thank T. Vlught for sharing his MC simulation results before the publication. N. L. Acknowledges the financial support by the Schwerpunktprogramm Polyelectrolyte of the German Science Foundation, DFG.

- [1] Rief M., Oesterhelt F., Heymann B. and Gau H. E., *Science*, **275** (1997) 1295.
- [2] Kellermayer M., Smith S. B., Granzier H. and Bustamante C., *Science*, **276** (1997) 1112.
- [3] Senden T. J., diMeglio J. and Auroy P., *Eur. Phys. J. B*, **3** (1998) 211.
- [4] Li H. B., Rief M., Oesterhelt F. and Gaub H. E., *Advanced Materials*, **10** (1998) 316.
- [5] Courvoisier A., Isel F., François J. and Maaloum J., *Langmuir*, **14** (1998) 3727.
- [6] Liphardt J., Onoa B., Smith S. B., Tinoco Jr. I. and Bustamante C., *Science*, **292** (2001) 733.
- [7] Borisov O. V. and Halperin A., *Eur. Phys. J. B.*, **9** (1999) 251.
- [8] Lee N., Johner A. and Vigils T. A., *preprint*.
- [9] Klimov D. K. and Thirumalai D., *Proc. Natl. Acad. Sci. USA*, **96** (1999) 6166.
- [10] Ramanathan S. and Shakhnovich E. I., *Phys. Rev. E.*, **50** (1994) 1303.
- [11] Doi M. and Edwards S. F., *The theory of polymer dynamics* (Oxford Clarendon Press, Oxford) 1986.
- [12] Jösön B., Peterson C. and Söderberg B., *Phys. Rev. Lett.*, **71** (1993) 376.
- [13] Migliorini G., Rostiashvili V. and Vilgis T. A., *Eur. Phys. J. E*, **4** (2001) 475.
- [14] Migliorini G., Lee N., Rostiashvili V. and Vilgis T. A., *Eur. Phys. J. E*.
- [15] Lee N. and Vigils T. A., *preprint*.
- [16] Feynman R. P., *Statistical Mechanics* (Addison Wesley, Redwood city) 1972.
- [17] Halperin A. and Zhulina E. B., *Euro. Phys. Lett.*, **15** (1991) 417.
- [18] Lai P.-Y., *Phys. Rev. E.*, **58** (1998) 6222.
- [19] Furukawa K. and Ebata K., *Appl. Phys. Lett.*, **77** (2000) 4289; Furukawa K., Ebata K. and Matsumoto N., *Appl. Phys. Lett.*, **75** (1999) 781.
- [20] Obukhov S. P., *J. Phys. A*, **19** (1986) 3655.
- [21] Sfatos C., Gutin A. M. and Shakhnovich E. I., *Phys. Rev. E.*, **48** (1993) 465.
- [22] Garel T. and Orland H., *Euro. Phys. Lett.*, **6** (1988) 597.
- [23] Vilgis T. A., Johner A. and Joanny J.-F., *Eur. Phys. J. E*, **2** (2000) 289.
- [24] Shakhnovich E. I. and Gutin A. M., *J. Phys. (France)*, **50** (1989) 1843.
- [25] Vlught T., private communication.